Data Protection of Internet Enterprise Platforms in the Era of Big Data

Jiaxing Zhang[1], Anuo Yang[1] and Feng Shuaishuai[2,*]

[1]*School of Social and Behavioral Sciences, Nanjing University, Nanjing, Jiangsu Province, China*
[2]*School of Sociology, Wuhan University, Wuhan, Hubei, China*
E-mail: dg21070024@smail.nju.edu.cn
[*]*Corresponding Author*

Received 30 November 2021; Accepted 27 December 2021;
Publication 08 March 2022

Abstract

With the development of big data technology, processed data has become an important source of value. Data has played a pivotal role in the development of enterprises, especially internet enterprises. However, Internet enterprise platform companies generally infringe on personal privacy in various stages of information collection, processing and application, and Internet enterprise platform data protection research is of great significance. The study found that the current problems of data protection on Internet enterprise platforms include: extremely weak user data protection measures, intellectual property risks throughout the whole process of big data processing, and infringements that have both new and high-tech characteristics. The high ambiguity in the definition and attribution of "data rights", the low cost and high concealment of infringements, and the value difference between intellectual property protection and digital economy are the main causes of these problems. As far as the protection path is concerned, we should start from the three aspects of technology empowerment, governance empowerment and legal

Journal of Web Engineering, Vol. 21_3, 861–878.
doi: 10.13052/jwe1540-9589.21314

empowerment, and work together to promote the proper protection of Internet enterprise platform data.

Keywords: Internet enterprise platform, data protection, cause analysis, protection path.

1 Introduction

The rapid development of science and technology such as Big Data, Cloud Computing, Internet of things (IOT) and Artificial Intelligence (AI) makes it possible for the production and application of massive data. Under this background, data elements are regarded as the basic resources of the fourth industrial revolution. Countries all over the world have successively launched their development strategies of digital economy and smart country construction. In the pre big data era, static and single data are mainly used as scientific research tools and have no economic value, and there was no discussion on data ownership and data protection. Data can be processed as Public Goods, and data developers can mine and use it freely. However, with the development of big data technology, processed data has become an important source of value and has far-reaching significance for individuals, enterprises and governments. For example, by using data, individuals can enjoy intelligent and personalized services, enterprises can significantly improve their decision-making, innovation and operation level, and governments can strengthen the monitoring of economic growth and population mobility. With the increase of data accumulation, every link from data collection and processing to data analysis and trading will produce huge market benefits (Shi, 2018). So far, data plays an important role in the development of enterprises, especially the internet enterprises. The more users an internet enterprise and platform has, the more users it may attract, and the more likely it is to be in a favorable position in the competition with other internet enterprises and platforms. This snowballing network effect makes internet enterprises often regard data as their core asset in the competition (Katz & Shapiro, 1985; McGowan, 1998). Nevertheless, just like a coin having two sides, with the ushering of the big data era, the digital economy has not only injected the lasting vitality into social and economic development, but also got mankind trapped and disturbed under the impact of the torrent of technology. In 2009, Professor Victor of Oxford University put forward that “the internet has put mankind into a digital circular prison”, in his book *Delete: The Virtue of Forgetting in the Digital Age*. Although the laws and regulations issued by countries all over the world have relevant provisions on the use and protection of Internet enterprise platform data, Internet enterprise platform data infringement and violations are still emerging one after another and have the characteristics of diverse forms and clever means. In order to promote the innovative application and healthy development of big data in the internet industry, it is necessary to analyze the problems faced by the current Internet enterprise platform data protection and their underlying causes.

2 Data Protection of Internet Enterprise Platforms: Current Situation and Problems

2.1 Current Situation of Data Protection of Internet Enterprise Platforms

The development of new communication technology and the internet has greatly enhanced the ability of domestic and international information exchange, which has a significant impact on the relationship between people and society and interpersonal relationship. At the same time, the rapid development of technology makes the collection and processing of personal information more common and universal, which poses a potential threat to individual rights and even impacts on those of the basics (Solove, 2006; Helbing and Stefano, 2011). Due to the various forms required to be filled in when surfing the internet, a series of activities such as the time spent on each web page and the columns clicked will be recorded in details. Coupled with the application of data analysis and data mining tools, it has become easy to obtain personal data and peep at people's online activities and even people's psychological activities through the internet. As Mantelero (2013) put it, "data is not only about memory, but also about power." In the era of big data, people live under digital surveillance beyond their awareness, they have no idea who is monitoring or what information is peeked by others; people simply become transparent to the public (Szekely, 2012). As a result, historical information has become a horrible existence, and anything you say may be used against you in the future (Costa and Poullet, 2012).

It is widespread that internet enterprises platform violate personal privacy in various stages of information collection, processing and application, which may be even more serious in some places. The main manifestations include: (1) excessive collection and unspecified use of personal data. Especially, it is often encountered that when registering a website, one needs to fill in a detailed personal data form including real name, gender, date of birth, valid certificate type and number, detailed mailing address and postal code, telephone number, personal home page, email, internet access means, weekly internet access time, province, city, industry, occupation and position, enterprise type and scale, highest education, income level and working conditions, etc. Most of the content is often not directly related to the purpose of using the website, not to mention who will use the data and for what purposes. (2) Illegal data transactions. Online customer information can be a valuable asset, some website owners under difficult operation conditions may sell it for profit. (3) Overexploitation, which refers to that some Internet enterprise platforms obtain depth data after analyzing and sorting or data mining from the personal data collected and use it for promoting their own business or other purposes, such as algorithmic recommendation, targeted marketing, or even online fraud.

Although countries around the world have successively issued a series of laws and regulations aimed at promoting data protection on the internet, such as The General Data Protection Regulation (GDPR) of the European Union, The Consumer Privacy Bill of Rights Act of 2015 (CPBR) of the United States, The Personal Information Protection Law of Japan (2020 Amendment) and The General Number Act, and The E-Commerce Law, The Data Security Law and The Personal Information Protection Law of China, but generally speaking, the current situation of data protection on internet enterprises platforms is not optimistic, and there still exist major problems.

2.2 Problems in Data Protection of Internet Enterprise Platforms

First, the user data protection measures currently used by internet enterprises platforms are extremely weak. Big data technology can discover hidden relationship patterns, resulting in the risk of personal information leakage such as identity information, location information and relationship information. In terms of the privacy policies provided by the current internet enterprises platforms, they are not enough to effectively protect users' privacy. One is that users do not have the opportunity to negotiate privacy policies with the platforms. The current practice is that users who register or use the services provided by the Internet enterprise platforms are deemed to have accepted the privacy policy of the platforms. For example, in its "Legal Statement and Privacy Policy", Taobao, a major Chinese e-commerce platform, pointed out that: "Once you start using Taobao's products or services, you have fully understood and agreed to the policy," which means that the privacy policy of the platform has been bundled with its services. Then, in terms of authorization scope, the platforms only tell users what kinds of data they collect. But usually, this is a general authorization, and the way the platforms use the information is very ambiguous. For example, China's Baidu search engine informs users through its "Privacy Protection Statement" that "user information is used to provide and improve products and services." But actually, the collected information is not only used for Baidu's own business and services, but also for Baidu's affiliates, partners, third-party suppliers, service providers and agents, etc. Finally, most current privacy policies lack data lifecycle rules and do not specify when to destroy the data. Once the platforms stop running or users no longer use them, the remaining large amount of user information will no longer be regulated, and it will become even more difficult to protect users' privacy.

Second, there exist intellectual property risks all over the links of the big data processing chain on Internet enterprise platforms. Big data processing at Internet enterprise platforms generally includes three links: data collection, storage, mining and analysis, which have varying degrees of intellectual property risks. One is the risk of infringement in data collection. Internet data collection is mainly carried out through crawler programs. The performance of a crawler directly affects the performance of the whole search engine, such as whether the content of the platform is rich and whether the information can be updated in time. Usually, the platform website will set crawler exclusion conditions to tell the search engine which pages can be crawled and which not. Some search engines will adopt anti-monitoring strategy (or anti-crawler) to simulate normal operation and realize continuous data capture. This violates the will of the website being visited and there is a risk of infringement. Then, the risk of infringement in data storage. Cloud computing platforms are for data storage, analysis and service. In computing, cache is a high-speed data storage layer, which usually temporarily stores a subset of data. The main purpose of caching is to reduce the access requirements to the slower underlying storage layer, so as to improve the data retrieval performance. However, such data storage services on the cloud computing platforms, especially the content copy or content cache, will face copyright infringement charges. Finally, the infringement risk in data mining and analysis. With the automation of data generation, the volume of data increases sharply. At the same time, each data is not isolated and static, but connected to each other. This means that even if the data has passed label processing, the state and behavior pattern of a specific user can still be identified based on behavior analysis and relationship analysis in a massive data environment. It can even accurately locate specific people, thus increasing the risk of personal information disclosure.

Third, infringement has the characteristics of new type, complexity and high technology. With the rapid development of big data technology, technological progress not only improves the ability of big data mining and analysis for enterprises, but also further increases the concealment of lawless infringement and illegal acts. There are many “Bag Companies” in the network, which engage in illegal activities such as stealing and selling personal privacy data in the name of digital economy. The black industrial chain that illegally obtains and divulges personal information has quietly emerged and has an obvious aggravating trend (Yan, 2018). For example, the data dispute of Huawei v. Tencent, the interface dispute of SF v. Cainiao Courier Station, the Sina v. Maimai case, the Dianping v. Baidu case, the Taobao v. Meijing unfair competition case, the Craigslist v. 3Taps case and the hiQ v. LinkedIn case that have emerged around the world in recent years are typical cases of illegal collection of platform data. The diversification of infringement forms and subjects, the expansion and digitization of infringed objects, the duality of the nature of infringement objects, the intellectualization and concealment of infringement means, and the seriousness, complexity and expansion of infringement consequences are all new problems brought to us in the era of big data and internet development.

3 Reasons for Problems in Data Protection of Internet Enterprise Platforms

3.1 Internal Driving Force: High Ambiguity in the Definition and Attribution of “Data Rights”

The ownership distribution of data rights on Internet enterprise platforms is an important internal reason for the weak data protection. Different types of data subjects hope to claim their exclusive rights to the data on the Internet enterprise platforms. So, should individuals, sensor equipment manufacturers, network service providers, software program manufacturers (investors) or other subjects enjoy the data alone, or should they enjoy the data jointly among the above different subjects, or should the data be directly classified into the public domain? This problem has directly led to the difficulty of defining the facts of data infringement and increased the difficulty of resorting to relevant laws to seek protection. Taking personal data as an example, the academic discussion on the ownership of personal data on Internet enterprise platforms can be summarized into four categories: first, personal ownership of the data. This means that once the personal ownership of data is regarded as an inalienable personality right, the collectors and users of the data shall not restrict the free exercise of this data right. Just as private individuals cannot restrict citizens' free use of personal names through contracts, enterprises cannot require individuals to give up their data rights through contracts (Hansmann and Kraakman, 2002). Second, Internet enterprise platforms own the data. As stipulated in the user agreement of Sina Weibo in the early days: "For the information released by the user on the microblog, including but not limited to words, pictures, videos, audio, etc., whether the microblog content constitutes a protected object in the sense of copyright law or not, the user agrees to irrevocably authorize the microblog platform as the exclusive publishing platform of the microblog content, and the microblog content published by the user will only be displayed exclusively on the microblog platform." Third, data is shared by individuals and Internet enterprise platforms. However, in the case of data sharing between individuals and Internet enterprise platforms, the division of power and the right boundary between individuals and the platforms are still a problem (Ding, 2019). Fourth, data is owned by the public domain. That is, once the platform is involved in the internet, it means that the platform data has a public attribute and is not owned by any private or enterprise. According to the above debate, it is very difficult to define the ownership of the data rights appropriately. It is at this level, Nissenbaumh (2004) proposed that the protection boundary of personal data is not rigid, but subjective and dynamic, which is affected by many factors, and the identification of personal data needs to be investigated in combination with specific scenarios.

3.2 External Driving Force: Low Cost Plus Strong Concealment Resulting in Frequent Infringements

Big data is known as "the oil in the digital age." It not only enables internet enterprises to update themselves continuously with the rapidly changing trend, but also has the ability to predict the future development trend, making internet enterprises to gain more competitive advantages. In this sense, data is increasingly becoming the core competitive resource of Internet enterprise platform operators. However, the cost of obtaining big data is very low. On the one hand, under the internet environment, players can successfully obtain data information by completing a series of operations such as domain name resolution, data resolution and programming calculation in a very short time.

And such operations only need very small investment. On the other, the internet market is known as the "invisible market" and all kinds of business operations and consumption behaviors are carried out on the internet. Many things are invisible and cannot be grasped, therefore, investigation and evidence collection are difficult due to the particularity. In addition to the technical and intangible fictionality of the internet market itself, the players may hide the traces of their data usage behaviors by relying on the technical processing, so that the behaviors are submerged in the vast data universe of the internet market, which is difficult to be found and regulated afterwards. Sokol and Comerford (2016) believe that after allowing data circulation, Internet enterprise platform enterprises form a data monopoly after collecting a large amount of data, which will undermine the normal market operation order and hinder the innovative development of the data industry. If the monopoly enterprises rely on the dominant position of data market to realize data rent-seeking, control data resources and grab additional benefits, it may greatly damage the interests of consumers.

3.3 Internal Conflict: The Value Difference Between Intellectual Property Protection and Digital Economy

In addition to the above two reasons, an internal contradiction of the data protection problems on Internet enterprise platforms lies in the value difference between what the big data emphasizing on "open sharing" and what the intellectual property emphasizing on "special protection". In practice, there is a formal conflict between data circulation and sharing and data protection. Data protection means to "close" the data in a specific field and prevent the demand side from obtaining it by setting thresholds and obstacles, so as to reduce the risk caused by flow. On the contrary, data sharing means that the data will be obtained and used by the subjects in need in a more open and diversified way, and the flow will be enhanced to improve the efficiency of data value-added (Chen and Gu, 2020). The opening and sharing mechanism of data is very important for the development of big digital economy, emphasizing the openness to public interests, but intellectual property rights are characterized by paying attention to the protection of private rights. Generally speaking, data generating subjects are usually unwilling to share or only selectively share unimportant data. Therefore, in the relationship between intellectual property and big data, there may be the conflict between the specificity of intellectual property and the sharing of data, the conflict between the regionality of intellectual property law and the infinity of data sharing, and the conflict between the timeliness of intellectual property rights and the rapid update speed of data. Therefore, this formal conflict naturally makes data sharing pose a severe challenge to the protection of user information, especially personal user information (Wang, 2019). Data protection should not become a legitimate excuse and legal cloak for data blockade, data monopoly and even abuse of data hegemony. From the characteristics of data itself, i.e., instantaneity, low-density and reusability of value, as well as the long-term needs of the development of digital economy, data must be safely and efficiently circulated and reused in order to better realize the value mining and innovation of data and truly promote the high-quality development of the digital economy.

4 Protection Path of Internet Enterprise Platform Data in the Era of Big Data

4.1 Technology Empowerment: AI Algorithm Helps Internet Enterprise Platform Data Protection

As pointed out above, the risk of data protection of Internet enterprise platforms exists in all links of data production, storage, analysis and mining and data use. Among them, the data protection of each link needs the corresponding algorithm technical support: (1) in the data production stage. The risk faced by big data production is how to efficiently and reliably remove the content that may leak users' privacy while ensuring the availability of users' data. In view of this, the anonymous publishing technology of big data, including k-anonymity (Sweeny, 2012), l-diversity anonymity (Barbaro & Zeller, 2006), t-close ness anonymity (Narayanan and Shmatikov, 2006), m-variance anonymity (Xiao and Tao, 2007), and anonymity based on "role composition" (Bu et al., 2008) etc. can realize the anonymity protection in the data production stage. (2) In the data storage stage. In the era of big data, the data storage provider is generally a cloud storage platform, and the user's data is faced with the risk of being peeped or tampered with by an untrusted third party. Accordingly, homomorphic encryption technology (Van et al., 2010), hybrid encryption technology (Chen and Huang, 2013), por model based on BLS short signature (Juels and Kaliski, 2007), DPDP (Erway et al., 2009), and Knox (Wang et al., 2012) etc. are feasible methods to prevent privacy disclosure during the data storage. (3) In the data mining and analysis stage. Privacy protection technology for data mining is to study more appropriate data hiding technology on the premise of improving the availability of big data as much as possible. At present, the main technologies include methods based on data distortion and encryption, such as data transformation, hiding, random disturbance, translation, inversion and other technologies (Agrawal and Philip, 2000; Oliveira and Zaiane, 2010). (4) In the data use phase. In order to solve the problem of privacy disclosure when accessing and using big data, the current technologies mainly include: spatio-temporal role based access control (Damiani et al., 2007), attribute-based encryption access control (ABE) (Goyal et al., 2006), ciphertext policy attribute set based encryption (CP-ASBE) (Bobby, Khurana and Prabhakaran, 2009), and hierarchical attribute set based encryption (HASBE) (Wan et al., 2012), etc.

In addition, the block chain technology plus artificial intelligence is also regarded as an effective technical means for data protection of Internet enterprise platforms in the big data era. From the perspective of technology, block chain is not a single technological innovation, but a distributed ledger technology realized after the deep integration of P2P network technology, asymmetric encryption technology, consensus mechanism, on-chain script and other technologies (Watanabe et al., 2016). The application of block chain technology is mainly reflected in the full life cycle management of Internet enterprise platform data, solving the problems existing in the confirmation, use, protection and transaction of digital content, and realizing the functions of digital copyright registration, intelligent transaction and infringement monitoring (Liu et al., 2018). Using the decentralized feature of block chain technology can increase the encryption of data transmission and reduce the risk of data leakage. As Lessig (1999) said, internet regulation cannot be limited to law, but needs the interaction of law, and social norms, market and technology, and technology can replace law and become an effective social governance tool in some fields.

4.2 Governance Empowerment: Coordinate the Whole Chain Protection of Internet Enterprise Platform Data

The data protection of Internet enterprise platforms in the big data era is a systematic project involving a wide range. It needs to comprehensively use multiple governance means such as administration, law, technology, media, economy and culture to improve the protection system and strengthen the coordination from multiple links such as review and authorization, administrative law enforcement, judicial protection, arbitration and mediation, public opinion advocacy, industry self-discipline and citizen integrity, and strive to build a whole chain of data protection work pattern. It should be noted that although there is a certain dilemma of improper protection of Internet enterprise platform data, this should not be the reason for the over strong and excessive intervention of public power in the development of digital economy in the big data era. The governance of the Internet enterprise platform data protection should be sparse rather than blocked, and the function of social subjects should be given full play. Among them, supporting and guiding the establishment of industry norms is particularly critical, specifically: (1) the data access system for enterprises to implement privacy protection mechanism should define three standards (Li and Cheng, 2012): one is the flexibility. Different people have their own privacy protection needs, so we should provide users with a flexible mechanism to set protection policies according to their needs. Second is the data quality. The quality of data should be guaranteed while protecting users' privacy. Third is the simplicity. The establishment of policies should be simple and easy to implement. (2) Comply with industry privacy laws. Some special industries will involve more complex privacy data management. Therefore, it is necessary to formulate more elaborate industry privacy laws to better protect personal privacy data. (3) Transfer control according to access rights. The data providers shall specify the purpose, conditions, retention time and responsibilities of data users accessing the data. We should also pay attention to the privacy level of the transmitted data, and use the combination of content encryption and auxiliary measures. (4) Build trust between enterprises and users. Patrick et al. (2005) stressed that an important factor in people's acceptance of the system is people's trust in the system. Therefore, in order to reduce users' concerns about their privacy, enterprises should try to establish an effective personal privacy data protection mechanism.

4.3 Legal Empowerment: Exploring Legal Potential and Realizing the Goal of Behavior Regulation

Considering the open characteristics of big data and digital economy, we believe that the strong protection mode of empowerment should not be adopted for the protection of Internet enterprise platform data, but a legal guarantee system containing multiple values should be constructed on the basis of making full use of existing laws, so as to achieve the purpose of behavior regulation of Internet enterprise platform data. The primary purpose of constructing a data sharing model with multiple values is to ease the tension and conflict between data flow sharing and privacy protection. Some people (Wang and Ye, 2019) believe that the emergence of tension and conflict between the two is due to the game and differences between public interests and private interests, data property interests and personality interests under the background of new technology. These differences can be bridged by resetting the rights allocation mode, that is, two stages of rights construction can be carried out on the basis of distinguishing personal information and data assets. The former configures personality rights and property rights, and the latter configures data management rights and asset rights (Long, 2017). Specifically, on the one hand, it is necessary to adopt a comprehensive and coordinated unified legislative model to give full play to the regulatory role of laws of various departments. The lack of independent intellectual property laws for big data on Internet enterprise platforms does not mean that it is not protected. Through relevant legal decisions, we can see that the copyright law, patent law, trademark law and anti-unfair competition law have begun to play a role. On the other, there should be different protection means for different types of data. First, personal information should be protected not only by general law, but also by special personal information protection law. Second, privacy data is mainly governed by privacy laws, trade secrets are usually protected by trade secret law, and data containing state secrets shall be protected by the national security law. Third, in order to trade personal data, anonymity/privacy processing is required. Original data and data controllers (i.e. data collected through capital investment) can sign data exchange agreements subject to contract law. Fourth, the data without personality rights attribute belongs to the legal controller, and the data involving public interests should be shared by the public for reasonable development and use. For example, there are points of views that it is necessary to distinguish personal data from non-personal data and believe that non-identifiable non-personal data does not involve the protection of the rights and interests of natural persons, and there is no need to give too many restrictions in law (Cheng, 2018). Fifth, data targeted at intellectual property rights are bound by intellectual property law.

Acknowledgements

This paper is Supported by the National Social Science Fund of China (16ZDA086, 20BSH002), and the Independent Research (Humanities and Social Sciences) at Wuhan University (413000057).

References

[1] Shi D. The Research on Data Property and Protection in Big Data Era[J]. Journal of Xi'an Jiaotong University (Social Sciences), 2018, 38(3): 78–85.

[2] Katz M and Shapiro C. Network Externalities, Competition, and Compatibility[J]. The American Economic Review, 1985, 75(3): 424–440.

[3] Mcgowan L D. Legal Implications of Network Economic Effects[J]. California Law Review, 1998, 86(3):479–611.

[4] Solove D J. Nothing to Hide: The False Tradeoff between Privacy and Security[J]. Social Science Electronic Publishing, 2006, 111(6): 1021–1043.

[5] Helbing D and Stefano B. Big Data, Privacy, and Trusted Web: What Needs to Be Done[M]. University Library of Munich, Germany, 2011.

[6] Mantelero A. The EU Proposal for a General Data Protection Regulation and the roots of the 'right to be forgotten'[J]. Computer Law & Security Review, 2013, 29(3):229–235.

[7] Szekely I. The Right to Forget, the Right to be Forgotten[C]. In: Gutwirth S, Leenes R, De Hert P & Poullet Y(eds). European Data Protection: In Good Health?[M]. Springer, Dordrecht, 2012, pp. 347–363.

[8] Costa L and Poullet Y. Privacy and the Regulation of 2012[J]. Computer Law and Security Review, 2012(28): 254–262.

[9] Yan C L. How to Prevent the "Streaking" of Personal Privacy in the Era of Big Data[J]. People's Tribune, 2018(16): 82–83.

[10] Hansmann H and Kraakman R. Property, Contract, and Verification: The Numerus Clausus Problem and the Divisibility of Rights[J]. Journal of Legal Studies, 2002, 31(S2):S373–S420.

[11] Ding X D. Who Owns the Data? Platform Data Ownership and Protection from the Perspective of Web Crawler[J]. ECUPL Journal, 2019, 22(5): 69–83.

[12] Nissenbaum H. Privacy As Contextual Integrity[J]. Washington Law Review, 2004, 79(1): 119–158.

[13] Sokol D D & Comerford R E. Antitrust and Regulating Big Data[J]. University of Florida Levin College of Law Research, 2016(SEP): 16–40.

[14] Chen B & Gu D D. Rethinking and Restructuring of the Rational Way of Data Sharing in Digital Economy: From the Perspective of

Data Typing[J]. Journal of Shanghai University of Finance and Economics(Philosophy and Social Science), 2020, 22(2): 122–137.

[15] Wang L M. Data Sharing and Personal Information Protection[J]. Modern Law Science, 2019, 41(1): 45–57.

[16] Sweeny L. K-anonymity: a Model for Protecting Privacy[J]. International Journal on Uncertainty, Fuzziness and Knowledge Based Systems, 2012, 10(5): 557–570.

[17] Barbaro M and Zeller T. A face is exposed for AOL searcher No. 4417749[N/OL]. New York Times, (2006-08-09). http://www.nytimes.com/2006/08/09/technology/09aol.html

[18] Narayanan A and Shmatikov V. How To Break Anonymity of the Netflix Prize Dataset[J]. Computer Science, 2006(Oct), arXiv:cs/0610105.

[19] Xiao X K and Tao Y F. M-invariance: towards Privacy Preserving Republication of Dynamic Datasets[C]. Proceedings of the 2007, ACM SIGMOD International Conference on Management of Data, June 12–14, 2007, Beijing, China. New York: ACM Press, 2007: 689–700.

[20] Bu Y Y, Fu A W C and Wong R C W, et al. Privacy Preserving Serial Data Publishing byRole Composition[C]. Proceedings of the 34th International Conference on Very Large Data Bases, August 23–28, 2008, Auckland, New Zealand. [S.l.: s.n.], 2008: 845–856.

[21] Van D M, Gentry C & Halevi S, et al. Fully Homomorphic Encryption over the Integers[C]. Proceedings of the 29th Annual International Conference on the Theory and Applications of Cryptographic Techniques, May 30–June 3, 2010, Riviera, French. New York: Springer Berlin Heidelberg, 2010: 24–43.

[22] Chen X and Huang Q. The Data Protection of Map Reduce Using Homomorphic Encryption[C]. Proceedings of the 4th IEEE International Conference on Software Engineering and Service Science (ICSESS), May 23–25, 2013, Beijing, China. Piscataway: IEEE Press, 2013: 419–421.

[23] Juels A & Kaliski B S. PORs: Proofs of Retrievability for Large Files[C]. Proceedings of the 14th ACM Conference on Computer and Communications Security, October 29-November 2, 2007, Alexandria, VA, USA. New York: ACM Press, 2007: 584–597.

[24] Erway C, KüPçü A and Papamanthou, et al. Dynamic Provable Data Possession[C]. Proceedings of the 16th ACM Conference on Computer and Communications Security, November 9–13, 2009, Chicago, IL, USA. New York: ACM Press, 2009: 213–222.

[25] Wang Q, Wang C and Li J, et al. Enabling Public Verifiability and Data Dynamics for Storage Security in Cloud Computing[C]. Proceedings of ESORICS, September 21–25, 2009, Saint Malo, France. [S.l.:s.n.], 2009: 355–370.

[26] Aggarwal C C and Philip S Y. A General Survey of Privacy-Preserving Data Mining Models and Algorithms[M]. New York: Springer US, 2008.

[27] Oliveira S M and Zaiane O R. Privacy Preserving Clustering by Data Transformation[J]. Journal of Information and Data Management, 2010, 1(1): 37.

[28] Damiani M L, Bertino E & Catania B, et al. Geo-rbac: A Spatially Aware rbac[J]. ACM Transactions on Information and System Security (TISSEC), 2007, 10(1): 2.

[29] Goyal V, Pandey O and Sahai A, et al. Attribute-based Encryption for Fine-grained Access Control of Encrypted Data[C]. Proceedings of the 13th ACM Conference on Computer and Communications Security, October 30–November 3, 2006, Alexandria, Virginia, USA. New York: ACM Press, 2006: 89–98.

[30] Bobba R, Khurana H and Prabhakaran M. Attribute-sets: A Practically Motivated Enhancement to Attribute-based Encryption[C]. Proceedings of the 14th European Symposium on Research in Computer Security, September 21–25, 2009, Saint-Malo, France. [S.l.: s.t.], 2009: 587–604.

[31] Wan Z, Liu J E and Deng R H. HASBE: A Hierarchical Attribute-based Solution for Flexible and Scalable Access Control in Cloud Computing[J]. IEEE Transactions on Information Forensics and Security, 2012, 7(2): 743–754.

[32] Watanabe H, Fujimura S and Nakadaira A, et al. Blockchain Contract: Securing a Blockchain Applied to Smart Contracts[C]. IEEE International Conference on Consumer Electronics. IEEE, 2016: 467–468.

[33] Li G J, Du X H and Wang N, et al. Research Progress of Blockchain Technology and Its Application in Information Security[J]. Journal of Software, 2018, 29(7): 2029–2115.

[34] Lessig L. The Law of the Horse: What Cyberlaw Might Teach[J]. Harvard Law Review, 1999, 113(2): 501–549.

[35] Li G J and Cheng X Q. Research Status and Scientific Thinking of Big Data[J]. Bulletin of Chinese Academy of Sciences, 2012, 27(6): 647–657.

[36] Patrick A, Marsh S and Briggs P. Designing Systems that People Will Trust[J]. Security and Usability, 2005, 1(1): 75–99.

[37] Wang Y and Ye M. Conflict and Balance between Personal Data Sharing and Privacy Protection in the Era of Artificial Intelligence[J]. Journal of Socialist Theory Guide, 2019(1): 99–106.

[38] Long W Q. On the Construction of New Data Property and its System Structure[J]. Tribune of Political Science and Law, 2017, 35(4): 63–77.

[39] Cheng X. Personal Data Rights in the Era of Big Data[J]. Social Sciences in China, 2018, 40(3): 174–188.

Biographies

Jiaxing Zhang is a PhD candidate in sociology at Nanjing University. Her research interests include Computational Social Science, Blockchain Technology and Social Governance, Big Data, etc.

She attended the Wuhan University where she received her B.Sc. in Software Engineering in 2009. Jiaxing Zhang then went on to pursuit a M.Sc. in software Engineering from Wuhan University, China in 2011. After that, she got a M.Sc. in Digital Media from Wuhan University, China in 2013.

Jiaxing Zhang has held solution and software engineering senior positions at Shenzhen since 2014. And she got some awards from some other research institutes in her research areas.

Anuo Yang is a PhD candidate in sociology at Nanjing University. Her research interests include the labor research, population migration, social class, and computational social science. She has published research include "The Impact of Perceived Discrimination, Positive Aspects of Caregiving on Depression among Caregivers: Mediating Effect of Job Satisfaction"" Prediction of Epidemic Spread of the 2019 Novel Coronavirus Driven by Spring Festival Transportation in China: A Population-Based Study", etc.

Feng Shuaishuai is a PhD candidate in sociology at Wuhan University. He received his bachelor's degree and master's degree in sociology from Northwest A&F University and Wuhan University respectively. His current focus is on computational social science research.